

Verfahren zur Verarbeitung von Postsachen

Beschreibung

5

Die Erfindung betrifft ein Verfahren gemäß dem Oberbegriff der unabhängigen Patentansprüche. Im weitesten Sinne bezieht sich die Erfindung auf das Gebiet automatischer Briefverarbeitung und insbesondere auf Systeme, bei denen ein automatisches Adreßleseverfahren durch Videocodierung bei der
10 Adreßinterpretation ergänzt und verbessert wird.

Systeme zum automatischen Lesen von Adressen (OCR) sind auf dem Gebiet der Briefbearbeitung wohlbekannt und z.B. in der
15 DE 195 31 392 beschrieben. Mit modernen OCR-Briefsortieranlagen sind Bearbeitungsraten von 10 Briefen pro Sekunde, d.h. 36.000 Briefen pro Stunde und mehr erreichbar. Die Erkennungszuverlässigkeit variiert jedoch stark mit der Schriftart und Gesamtqualität der auf die Oberfläche der Briefe aufgebrachten Adreßinformation. Im Falle einer erfolgreichen Erkennung kann der betreffende Brief mit einem maschinenlesbaren Barcode versehen werden. Dieser Barcode erlaubt eine weitere mechanische Bearbeitung bis zu einer beliebigen gewünschten Sortierordnung. Insbesondere ermöglicht
20 die Verwendung von Barcodes eine Sortierung der Briefe bis zur Sortierebene des Postgangs, bei der Briefe entsprechend der Reihenfolge ihrer Verteilung durch den Zusteller sortiert werden.

30 Da die Erkennungsraten der automatischen Lesesysteme sehr variieren, ist es notwendig, diese durch verschiedene Formen von manueller Intervention zu unterstützen. Die einfachste Intervention ist Zurückweisung nicht automatisch lesbarer Briefe, und die Durchführung eines Handsortierverfahrens. Allerdings sind die dadurch entstehenden Kosten mit zunehmenden
35 Arbeitskosten unökonomisch hoch. Hinzu kommt, daß solche handsortierte Post auch zu späteren Zeitpunkten nicht mehr

ohne weiteres mechanisch sortiert werden kann, so daß zwei voneinander getrennte Ströme von Sendungen erzeugt werden, die wiederum manuell zu einem bestimmten Zeitpunkt verschmolzen werden müssen.

5

Um diese Nachteile manueller Sortierung von OCR-zurückgewiesenen Sendungen zu vermeiden, sind verschiedene Verfahren für eine manuelle Codierung von Poststücken entwickelt worden. Alle diese Verfahren verwenden Interventionen durch Operatoren, um Barcodes auf die Sendungen auf eine Weise aufzubringen, die konsistent ist mit dem Erfordernis, eine mechanische Sortierung durch die gleichen Maschinen vorzunehmen, die OCR-gelesene und barcodierte Post bearbeiten.

15

Ein weiteres Verfahren um zurückgewiesene Poststücke zu codieren verwendet sogenannte manuelle Codierplätze. An diesen manuellen Codierplätzen werden die Sendungen physikalisch hintereinander einem Operator vorgeführt, wobei der Operator für jedes dieser Sendungen so viel Information codiert wie notwendig ist, um den Bestimmungsort eindeutig zu identifizieren. Dabei wird die eingegebene Adresse mittels eines Verzeichnisses in einen Sortierbarcode umgewandelt, der auf die Sendung aufgebracht wird. Die codierten Sendungen werden dann mittels Barcodesortierer (BCS), die identisch mit OCR-geeigneten BCS sind, weiterbearbeitet. Solche manuellen Codierplätze wurden als erste von der US-Post und der Royal Mail in den 70er Jahren eingeführt. Die Hauptnachteile solcher Vorrichtungen bestehen in der notwendigen Entfernung von Sendungen aus dem OCR-Sendungsstrom und ergonomischen Schwierigkeiten für den Operator beim Erkennen der ihm vorbeigeführten Sendungen.

20
25
30

Der nächste Fortschritt in der Behandlung von OCR-zurückgewiesener Sendungen war die Entwicklung von On-line-Video-codiersystemen (OVS). In einem OVS wird ein Videobild der Sendung einem Operator zum Codieren vorgeführt, statt der physikalischen Sendung bei den manuellen Codierplätzen. Das

35

Videobild wird dem Operator gezeigt, während die physikalische Sendung in Verzögerungsstrecken gehalten wird. In diesen Verzögerungsstrecken wird normalerweise die Sendung in Bewegung gehalten für eine Zeitdauer, die ausreicht, damit der OVS-Operator die notwendige Sortierinformation für das betreffende Bild eingibt. Die üblichen Verzögerungsstrecken erlauben eine Verzögerung zwischen 10 und 30 Sekunden. Je länger die Verzögerungsstrecke, desto größer sind die Kosten sowie die Erfordernisse für Wartung und die physische Größe der Anlage.

Das Hauptproblem bei der Verwendung von OVS ist, daß die zur Verfügung stehende Zeit lediglich ausreicht für eine sorgfältige Eingabe der Postleitzahl (ZIP) oder des Postcodes (PC), es sei denn, daß unpraktikabel lange Verzögerungsstrecken verwendet werden.

Daher wurden spezielle Codierverfahren entwickelt, um die notwendige On-line-Verzögerungszeit möglichst gering zu halten.

Um die Codierproduktivität zu erhöhen und/oder die Angabe von sämtlichen Adreßelementen, d. h. ZIP/PC, Straße/Postfach, Adressat/Postfach, Adressat/Firma, zu ermöglichen, sind im Stand der Technik verschiedene Methoden entwickelt worden. Im wesentlichen sind dies:

Vorschau-Codierung

Bei der Vorschau-Codierung erfolgt ein simultaner Display der Bilder von zwei Sendungen erfolgt; eines über dem anderen. Hierbei ist das niedrigere Bild das aktive, d. h. dessen Daten codiert werden. Nach einem geeigneten Training ist es Operatoren möglich die Information auf dem unteren Bild zu codieren, während sie bereits die Adreßinformation vom oberen Bild aufnehmen. Das obere Bild wird anschließend aktiv und der Prozeß wird fortgesetzt. Mit der Vorschau-Codierung ist es möglich, durch eine vollständige Überlappung der kogniti-

ven und motorischen Funktionen beim Codieren aufeinanderfolgender Abbilder die Operator-Produktivität zu verdoppeln.

Extraktionscodierung

- 5 Da bei dem praktisch erreichbaren On-line-Verzögerungs-Zeiten lediglich die ZIP/PC Adreßelemente verläßlich vom Operator eingegeben werden können, werden beim Extraktionscodieren bestimmte Schlüsselbestandteile der sich auf die Straße beziehenden Adreßbestandteils eingegeben. Üblicherweise basiert
10 die Extraktionscodierung auf speziell entwickelten Regeln, bei denen ein Code fester Länge, als Zugangsschlüssel zu einem Adreßverzeichnis verwendet wird. Zum Beispiel verwendet die Royal Mail eine Extraktionsformel, die auf den ersten drei und den letzten zwei Buchstaben basiert. Dabei müssen
15 Sonderregeln vom Operator auswendig gelernt werden, um überflüssige Adreßinformationen zu vermeiden und bestimmte unterscheidende Merkmale wie z.B. Richtungen zu berücksichtigen, z.B. Ost, West oder Kategorien z.B. Street, Lane, Road.
- 20 Extraktionscodierung besitzt trotz ihrer gewissen Effektivität einige größere Nachteile; insbesondere komplexe Extraktionsregeln, die häufig die Berücksichtigung des Endes eines Straßennamens erfordern, während diese Bestandteile üblicherweise am unklarsten geschrieben sind. Außerdem findet man ei-
25 ne signifikant hohe Rate von nicht eindeutigen Extraktionen bei denen mehrere Einträge in einem Verzeichnis dem Extraktionscode entsprechen, so daß keine eindeutige Sortierentscheidung getroffen werden kann. Desweiteren ist zu berücksichtigen, daß die Eingabeproduktivität der Operatoren vermindert
30 wird, sobald vom Operator Entscheidungen getroffen werden müssen, statt einer einfachen repetitiven Tastatureingabe.

Komplettierungscodierung

- Bei der Komplettierungscodierung wird im Kontrast zur Extraktionscodierung für jede zu kodierende Adresse eine variable
35 Eingabe vorgenommen. Im wesentlichen wird bei der Adreßeingabe ein Abgleich mit dem Adreßverzeichnis vorgenommen, bis ei-

ne eindeutige Übereinstimmung erreicht ist. Durch die Darstellung des Rests der Adresse sobald eine eindeutige Teilübereinstimmung identifiziert worden ist, wird ein Beschleunigungseffekt erreicht. Probleme treten bei dieser Technik jedoch dadurch auf, daß dem Operator ein Eingabestop-Signal zugeführt werden muß und eine Darstellung des identifizierten Adreßrests notwendig ist, was zu einer verminderten Eingabeproduktivität führt und eine Vorausschau-Codierung unmöglich macht.

10

Operatorassistierte OCR-Technik

Zur Erhöhung der On-line zur verarbeitenden Adreßinformation hat die US Post mit Operator assistierten OCR-Techniken experimentiert. Hierbei wird zur Erhöhung der Effektivität der Teil des Adreßbildes hervorgehoben, bei dem die OCR-Erkennung versagt hat. Da die Operatoren bei der Entzifferung von fehlenden Buchstaben langsam sind und zum Teil auch komplexe Erkennungsfehler, z.B. Segmentierungsprobleme auftreten, ist bei diesem Verfahren die Operatorproduktivität häufig niedriger als bei einfachen Wiedereingabe der betreffenden Adresse.

20

Off-line Codierung

Da bei keiner der oben erwähnten Codierungstechniken eine ausreichend hohe Produktivität bei reiner On-line-Codierung erreicht wird, wurde kürzlich ein off-line Codiersystem eingeführt, wie in der US PS 49 92 649 beschrieben. Bei diesem System werden Sendungen mit nicht erkannten Adressen mit einer zusätzlichen Information, einer tracking identification (TID) versehen. Die nicht erkannten Sendungen werden extern gespeichert, während die Bilder dieser Sendungen Operatoren zum Codieren präsentiert werden wobei keine zeitlichen Einschränkungen vorliegen. Anschließend werden die Sendungen TID-Lesegeräten vorgeführt. Die TID wird mit der eingegebenen Adreßinformation verknüpft. Darauf basierend kann auch eine übliche Barcode-Sortierinformation auf die Sendung aufgebracht werden, so daß die betreffende Sendung wie üblicherweise OCR-gelesene Sendungen bearbeitet werden können. Obwohl

35

das off-line-Video-Codierverfahren eine effektive Methode zur Codierung sämtlicher Adreßbestandteile darstellt, werden zusätzlich Kapazitäten für die Weiterverarbeitung mit nichtgelesenen Adressen versehenen Sendungen und eine entsprechend
5 komplexe Logistik verlangt.

Aus der EP-A-589119 ist ein Verfahren zum Finden der Adreßblöcke in den Abbildungen der Sendungsoberflächen beschrieben. Sind diese bei der automatischen Suche nicht eindeutig
10 lokalisiert worden, so werden die Abbilder anschließend einer Videocodierung zugeführt. Dabei kennzeichnet der Operator dieses Gebiet, das dann gespeichert wird. Dann wird die automatische Adreßblocksuche damit belehrt, so daß die Adreßblöcke für weitere gleichgeartete Sendungen von Großkunden auto-
15 matisch gefunden werden können. Dieses Verfahren betrifft eine der Adreßinterpretation vorgelagerte Stufe.

Aufgabe der vorliegenden Erfindung ist es, in möglichst kurzer Zeit die Adressen von Sendungen erfolgreich in einem
20 Adreßlesesystem mit OCR-Einheit und Videocodiereinheit zu lesen.

Die Erfindung ermöglicht eine verbesserte Integration von automatischen Lesesystemen und Videocodierung. Weiter ermöglicht die Erfindung die effektive Verwendung einer Extraktionscodierung in integrierten automatischen und Videocodiersystem, insbesondere auch eine Vereinfachung des Entscheidungsproblems für die Operatoren bei der Adreßcodierung. Ein weiterer Vorteil des erfindungsgemäßen Verfahrens ist, daß zusätzliche Sortierinformationen, z.B. solche, die sich auf die
30 Namenszeile der Adresse beziehen, effektiv ausgewertet werden können. Ein weiterer Vorteil der Erfindung besteht darin, daß Inkonsistenzen, die sich z.B. bei der Extraktionscodierung oder durch Fehler bei der Beschriftung von seitens des Absenders ergeben, effektiv behandelt werden können. Das Verfahren
35 ermöglicht weiter eine einfache Integration von on-line und off-line-Codierung sowie des Vorschau-Codierungsverfahrens.

Vorteilhafte Ausführungsformen der Erfindung sind den abhängigen Ansprüchen und der Beschreibung zu entnehmen.

- 5 Im folgenden wird die Erfindung anhand von Zeichnungen beispielhaft dargestellt. Im einzelnen zeigen:

FIG 1 eine schematische Darstellung einer Vorrichtung zur Ausführung des Verfahrens

10 FIG 2 und 3 eine Übersicht über den Datenfluß gemäß der Erfindung

FIG 1 zeigt eine schematische Darstellung einer Briefverteilanlage mit der das erfindungsgemäße Verfahren ausgeführt werden kann. Der OCR-Briefsortierer 100 besteht aus einer Zuführeinrichtung 110, die sukzessive Sendungen aus einem Magazin 111 abzieht und mit ca. 10 Sendungen pro Sekunde zu einem hochauflösenden Videoscanner 120 transportiert. Anschließend werden die Sendungen in einer Verzögerungsstrecke 121 transportiert. Die Sendungen weisen üblicherweise auf ihrer Oberfläche Adreßinformationen auf. Im OCR-Prozessor 130 erfolgt eine Auswertung der Adreßinformationen der Abbilder der Sendungen die vom Videoscanner 120 gewonnen wurden. Im Fall einer vollständigen Auswertung wird ein Barcodedrucker 150 angesteuert und die Sendung wird mit einem entsprechenden Barcode für die anschließende Sortierung in Sortierfächer 160 versehen. Der OCR-Prozessor 130 besteht aus einem oder mehreren Mikroprozessoren 131 mit assoziiertem Speicher 132 um Abbilder der Sendungen zu speichern. Weiterhin beinhaltet der OCR-Prozessor ein Adreßverzeichnis 134 mit ZIP-Codes, Städtenamen und Straßennamen und evtl. weiterer adreßbezogener Information. Der bei der Auswertung der Adreßinformation aufweisenden Abbilder erfolgt eine merkmalsgesteuerte Reduktion der aus dem Adreßverzeichnis gewonnenen Eintragung, so daß eine Art Teilwörterbuch erzeugt wird. In einzelnen Eintragungen werden dabei Glaubwürdigkeiten zugeordnet, so daß bei der Auswertung eine Anzahl von Daten von richtig erkannten Adressen erzeugt werden. Die Vorrichtung enthält weiterhin einen

Imagecontroller 170 sowie eine Anzahl Video-Codierplätze 200, die mit dem Imagecontroller 170 direkt oder durch ein lokales Netzwerk (LAN) 171 verbunden sind. Falls die OCR-Auswertung eines Abbildes nicht, d.h. vollständig erfolgreich war, wird
5 dieses Bild vom OCR-Prozessor 130 zum Image-Controller 170 transferiert, der einerseits den TID Barcodedrucker 151 steuert und andererseits das entsprechende Bild zu einer der Videocodierplätze 200 sendet. Der TID Barcodeprinter 151 bringt auf die entsprechende Sendung einen Identifikationscode TID
10 auf, der es zu einem späteren Zeitpunkt ermöglicht, die ausgewertete Adreßinformation mit der physikalischen Sendung zu knüpfen. Die Auswertung der Abbilder erfolgt in diesem Fall vorzugsweise off-line, obwohl grundsätzlich bei einer genügend langen Verzögerungszeit auch eine on-line-Auswertung
15 durch Videocodierung möglich ist. Im letzteren Fall kann der TID auch zu einem späteren Zeitpunkt, d.h. wenn die Videocodierung innerhalb einer bestimmten vorgegebenen Zeit nicht zu einer vollständigen Auswertung geführt hat, auf die Sendungen aufgebracht werden.

20

Zur Ausführung des weiteren genauer dargestellten Verfahrens ist der Image-Controller 170 derart ausgebildet, daß von der Videocodierung nicht vollständig ausgewertete Adreßinformationen unter Verwendung der Ergebnisse der Videocodierung im
25 OCR-Prozessor einer weiteren automatischen Adreßauswertung zugeführt werden.

FIG 2 und 3 zeigen Datenflußdiagramme. Vorzugsweise arbeiten die Operatoren mit geteilten Displays 210 von denen der obere
30 eine Vorschau ermöglicht und der untere üblicherweise aktiv ist. Vom Operator eingegebene Daten werden in der Promptzeile 211 dargestellt. Zum Videocodieren gelangen solche Abbilder von Sendungen, die nicht vollständig automatisch ausgewertet wurden 220. Beim Beispiel der FIG 2 erfolgt eine Eingabe eines ZIP-Codes „4431“, eines Extraktionscodes für die Straße
35 „Hell“ sowie die Hausnummer „8“. Statt eines zweigeteilten Displays sind auch andere Formen möglich. Diese Eingabeinfor-

mationen werden benutzt, um im Adressenverzeichnis 134 passende Einträge zu identifizieren. Für den Fall, daß eine eindeutige Zuordnung zwischen der eingegebenen Information und einem Eintrag im Adressenverzeichnis 134 gefunden wurde, ist damit grundsätzlich eine vollständige Auswertung der Adreßinformation des betreffenden Abbildes erfolgt. Allerdings ist für einen gewissen Prozentsatz von Eingaben eine solche Eindeutigkeit nicht erreichbar, da eine Anzahl verschiedener Adreßeinträge der codierten Adreßinformation entsprechen.

Gemäß FIG 3 erfolgt eine Entscheidung 300, ob eine vollständige Auswertung der Adreßinformation eines Abbildes bei der Videocodierung erfolgt ist. Ist die Entscheidung positiv (Yes) kann die entsprechende Sendung entweder mit einem Barcode versehen werden, wenn die Verzögerungszeit ausgereicht hat, um eine On-line-Video-Codierung durchzuführen oder es wird eine entsprechende Verknüpfung der TID und eine darauf basierende Barcodierung vorgenommen. In jedem Fall kann die entsprechende Sendung mit üblichen Mitteln weiter sortiert werden. Bei einer negativen Entscheidung (No) erfolgt erfindungsgemäß eine weitere automatische Auswertung unter Verwendung der Ergebnisse der Videocodierung, d.h. bei dieser weiteren automatischen Auswertung stehen dem OCR-Prozessor neben den auf der Abbild dargestellten Informationen die durch die Videocodierung gewonnenen Informationen zur Verfügung; im obigen Beispiel das Tripel „4432“, „Hell“, „8“. Dies ist in FIG 3 symbolisch durch den Inhalt des Kreises 310 zum Ausdruck gebracht. Anschließend erfolgt eine Entscheidung 320, ob nunmehr eine vollständige Auswertung des betreffenden Bildes stattgefunden hat. Im positiven Fall (Yes) wird die entsprechende Information zum weiteren Sortieren der Sendung verwendet analog wie nach einer positiven Entscheidung beim Entscheidungspunkt 300. Ist die Entscheidung negativ (No) erfolgt eine weitere Videocodierung unter Verwendung der Ergebnisse der weiteren automatischen Auswertung. Vorzugsweise werden dabei dem Operator eine Anzahl von Alternativen zur

Selektion vorgeführt, aus denen eine Selektion vorzunehmen ist.

5 Vorzugsweise werden die beiden letzten Schritte des Verfahrens nämlich die weitere automatische Auswertung sowie die weitere Videocodierung nicht mehr on-line sondern off-line erfolgen, weil die zur Verfügung stehende Verzögerungszeit zu gering ist für eine on-line-Ausführung.

10 Das erfindungsgemäße Verfahren beinhaltet demnach drei zusammenwirkende Phasen. Diese sind:

1. Dateneingabephase durch Videocodierung, bei der vorzugsweise mit einem einfachen Extraktionscode eine Codierung bestimmter Teile der Adreßinformation erfolgt. Üblicherweise wird hierbei die ZIP/PC Information sowie ein größerer Teil der weiteren Adreßinformation vollständig ausgewertet oder es erfolgt bei fehlenden ZIP/PC eine Extraktion des Ortsnamens. Dieser Phase vorgeschaltet war bereits eine erste automatische Auswertung der Adreßinformation.

25 Vorzugsweise erfolgt eine Eingabe mit einem geteilten Display. Für Vereinfachung der Eingabe wird ein einfacher Extraktionscode verwendet, z.B. ein 4stelliger Postcode, die ersten vier Alpha-Zeichen des Straßennamens und die Ziffern der Hausnummern der betreffenden Adresse. Bei dieser Extraktionscodierung ist eine Anpassung an die jeweiligen Postkonventionen ohne weiteres möglich. So kann z.B. die Anzahl der ersten Buchstaben variiert werden. Vorzugsweise wird der Postcode vom Operator nur dann eingegeben, wenn die OCR-Auswertung überhaupt kein Ergebnis gebracht hat. Die Mehrzahl der Sendungen wird daher die Eingabe von Straßenangaben betreffen. Vorzugsweise kann auch eine Strukturierung der Video-Codierung derart erfolgen, daß eine Gruppe von Operatoren Postcode und Straßenangaben eingibt, während

eine andere Gruppe lediglich Straßenangaben eingibt. Da ein bestimmter Prozentsatz der Post heutzutage Postfachangaben enthält, wird vorzugsweise eine geeignete Taste mit einem Postfachschlüssel belegt, die gegebenenfalls vom Operator gedrückt wird, wobei anschließend die Eingabe der Postfachnummer folgt. Für Firmenadressen, bei denen Straßen- oder Postfachangabe fehlen, ist auch die Eingabe der Firmenzeile möglich.

2. Weitere automatische Auswertung unter Verwendung der Informationen, die bei der vorhergehenden Phase durch Videocodierung eingegeben worden sind. Die zusätzliche Information erhöht die Wahrscheinlichkeit einer vollständigen Auswertung bei der weiteren automatischen Auswertung.

Bei nicht vollständig in der vorhergehenden Phase ausgewertete Adreßinformationen folgt eine weitere automatische Auswertung. Wird mit einem Extraktionscode gearbeitet, so liegen für diese zwei oder mehr Eintragungen im Adreßverzeichnis vor. Bei geeigneter Extraktionscodierung sind dabei in über 90% der Fälle lediglich zwei Adreßeinträge vorhanden.

3. Weitere Videocodierung derjenigen Abbilder, für die bei den vorhergehenden Phasen keine vollständige Auswertung möglich war. Vorzugsweise werden Abbilder der nicht ausgewerteten Adressen dem Operator präsentiert, zusammen mit den Ergebnissen der vorhergehenden Videocodierung und der automatischen Auswertung der zweiten Phase zwei. Vorzugsweise hat der Operator hierbei lediglich eine Selektion aus vorgegebenen Alternativen vorzunehmen.

Bei der weiteren automatischen Auswertung liegt nunmehr zusätzliche Kontextinformationen vor und zwar ist die Anzahl der zu berücksichtigen Adreßeinträge beschränkt durch die Anzahl der aus der Extraktionscodierung erhal-

tenen Kandidaten. Es kann davon ausgegangen werden, daß unter diesen Kandidaten die richtige Adresse zu finden ist. Zusätzlich ist üblicherweise die Hausnummer bekannt.

5

Bei der weiteren Videocodierung solcher Adreßinformationen, die in den vorhergehenden Phasen nicht vollständig ausgewertet wurden, werden vorzugsweise nicht eindeutige Ergebnisse der Extraktionscodierung oder zusätzliche
10 sortierrelevante Informationen, die sich auf der Adressatenlinie befinden, bearbeitet. Vorzugsweise werden dem Operator sukzessive Bilder der Sendungsoberfläche präsentiert, wobei in einem Fenster die Auswertungsalternativen der Adresse oder des Adressaten dargestellt sind.
15 Die Auswahl kann entweder durch Tastatureingabe einer Auswahlnummer, per Maus oder über einen Sprachprozessor erfolgen.

20

Bei einer bevorzugten Ausführungsform der Erfindung erfolgt eine Auswertung eines ersten Bestandteils der Adreßinformation und eine Auswertung eines zweiten Bestandteils der Adreßinformation und eine Prüfung der Ergebnisse dieser Auswertungen in Hinblick auf gegenseitige Konsistenz. Der erste Bestandteil der Adreßinformation kann insbesondere der ZIP-
25 oder Postcode sein, der zweite eine Straßenangabe oder eine Hausnummer. Inkonsistenzen zwischen diesen beiden Bestandteilen können z.B. durch Lesefehler oder durch eine Falschangabe des ZIP- oder Postcodes vorkommen. Bei einer Falschangabe werden ausgehend von den ersten drei oder vier Ziffern des
30 ZIP oder Postcodes eine Anzahl von Alternativen für den Straßennamen gewonnen. Vom Operator erfolgt eine Extraktionscodierung des zweiten Bestandteils der Adreßinformation, woraus ebenfalls eine Anzahl von Vorschlägen für den Straßennamen gewonnen wird. Bei der Konsistenzprüfung werden nur diejeni-
35 gen Vorschläge nicht zurückgewiesen, bei denen die Ergebnisse dieser beiden Auswertungen untereinander kompatibel sind.

Im folgenden wird anhand von zwei Beispielen die obengenannte Ausführungsform genauer erläutert. Auf einer Sendung sei die Adresse

5 Bucklestr. 5
 D-78457 Konstanz

statt der korrekten Adresse

10 Bücklestr. 5
 D-78467 Konstanz

aufgetragen. In diesem Fall wird bei einer Extraktionscodierung „78457 Buck 5“ verwendet um aus einem Straßenverzeichnis die Eintragung „Buckley 5, Konstanz“ zu erhalten. Dies ist eine korrekte Zuordnung der fehlerhaften Postleitzahl 78457. Bei der Konsistenzprüfung wird durch automatische Auswertung die Inkonsistenz von „Bucklestr.“ und „Buckley“ festgestellt und das betreffende Auswertungsergebnis zurückgewiesen. Wird andererseits die Adreßzeile mit dem ZIP-Code und der Ortsangabe mit einer hohen Fehlerrate OCR gelesen, während die Straßenangabe mit niedriger Fehlerrate gelesen wird, folgt vom Operator lediglich eine Eingabe des ZIP-Codes oder eine Anzahl von Buchstaben der Ortsangabe. Vorzugsweise wird in diesem Fall dem mit geringer Fehlerrate behafteten OCR-Ergebnis Vorrang gegenüber der Operatoreingabe gegeben.

Bei einer weiteren bevorzugten Ausführungsform der Erfindung erfolgt eine Auswertung der auf die Oberfläche von Sendungen aufgebrachten Angaben durch Videocodierung in den Fällen, in denen eine automatische Auswertung deshalb nicht erfolgreich war, weil Adreßangabe und Absenderangabe von der automatischen Auswertung nicht differenziert werden konnten. Dies tritt vor allem bei Post auf, bei der die Absenderangabe unmittelbar über oder unter der Adreßangabe aufgebracht ist, z. B. dänischer Post. Da beim Videocodieren eine Darstellung der Sendungsoberfläche für den Operator erfolgt, ist es üblicher-

weise leicht möglich, die betreffenden Informationen als
Adreß- bzw. Absenderangabe zu identifizieren und gegebenen-
falls eine Extraktionscodierung vorzunehmen.

Patentansprüche

1. Verfahren zur Bearbeitung von Sendungen in einem automatischen Adreßlesesystem, wobei ein Abbild der die Adreßinformation aufweisenden Oberfläche jeder Sendung gewonnen und einer
5 OCR-Einheit zur automatischen Auswertung zugeführt wird, und bei nicht eindeutig erkannter Adreßinformation das zugehörige Abbild an einen Videocodierplatz zur Videocodierung weitergeleitet wird, d a d u r c h g e k e n n z e i c h n e t, daß
10 das Abbild jeder innerhalb einer bestimmten Zeit mittels Videocodierung nicht eindeutig erkannten vollständigen Adreßinformation mit den bei der Videocodierung erhaltenen Informationen über erkannte Adreßteile der OCR-Einheit zu einer weiteren automatischen Auswertung zur Adreßinterpretation zugeführt wird.
15

2. Verfahren nach Anspruch 1, d a d u r c h g e k e n n z e i c h n e t, daß das Abbild jeder bei der weiteren automatischen OCR-Auswertung zur Adreßinterpretation nicht eindeutig erkannten vollständigen Adreßinformation mit den dabei
20 erhaltenen Informationen einem Videocodierplatz zur weiteren Videocodierung zugeführt wird.

3. Verfahren nach Anspruch 1 oder 2, d a d u r c h g e k e n n z e i c h n e t, daß bei der ersten Videocodierung eine Extraktionscodierung gemäß vorgegebener Extraktionsregeln erfolgt.
25

4. Verfahren nach Anspruch 2, d a d u r c h g e k e n n z e i c h n e t, daß bei der weiteren Videocodierung eine Selektionscodierung derart erfolgt, daß aus einer Anzahl von alternativen Auswertungsergebnissen eine Selektion vorgenommen wird.
30

5. Verfahren nach Anspruch 4, d a d u r c h g e k e n n -
z e i c h n e t, daß die alternativen Auswertungsergebnisse
aus den nicht eindeutigen Ergebnissen der Extraktionscodie-
rungen gebildet werden.

5

6. Verfahren nach Anspruch 5, d a d u r c h g e k e n n -
z e i c h n e t, daß alternative Auswertungsergebnisse aus
Adreßinformationen, die zusätzliche Sortierinformationen auf-
weisen, gebildet werden.

10

7. Verfahren nach einem der Ansprüche 1 bis 6, d a -
d u r c h g e k e n n z e i c h n e t, daß eine Auswertung
eines ersten Bestandteils der Adreßinformation erfolgt, und
daß eine Auswertung eines zweiten Bestandteils der Adreßin-
formation erfolgt, und daß die Ergebnisse dieser Auswertungen
in Hinblick auf gegenseitige Konsistenz geprüft werden.

15

8. Verfahren nach Anspruch 7, d a d u r c h g e k e n n -
z e i c h n e t, daß das Abbild jeder innerhalb einer be-
stimmten Zeit mittels Videocodierung nicht eindeutig erkann-
ten vollständigen Adreßinformation mit den bei der Videoco-
dierung erhaltenen Informationen über erkannte Adreßteile der
OCR-Einheit zu einer weiteren automatischen Auswertung zuge-
führt wird.

25

9. Verfahren nach Anspruch 1 bis 8, d a d u r c h g e -
k e n n z e i c h n e t, daß die erste automatische Auswer-
tung der Adreßinformation On-line oder Off-line erfolgt.

30

10. Verfahren nach Anspruch 1 bis 9, d a d u r c h g e -
k e n n z e i c h n e t, daß Sendungen, für die on-line keine
vollständige weitere automatische Auswertung oder Auswertung
durch weitere Videocodierung der Adreßinformation erfolgt
ist, mit einer Identifizierungsmarkierung (TID) für eine off-
line vorzunehmende weitere automatische oder Videocodierung
versehen werden.

35

11. Verfahren nach Anspruch 1 bis 9, d a d u r c h g e -
k e n n z e i c h n e t, daß zumindest bei einer der Videoco-
dierungen ein Vorschaucodierungsverfahren verwendet wird.

5 12. Verfahren nach Anspruch 1 bis 11, d a d u r c h g e -
k e n n z e i c h n e t, daß bei der Videocodierung eine Dif-
ferenzierung zwischen Adreß- und Absenderinformation erfolgt.

10 13. Vorrichtung zur Durchführung des Verfahrens nach Anspruch
1 bis 12, mit einem automatisches Adreßlesesystem, das
- eine Vorrichtung zur Gewinnung von Abbildern von Sendungen
(120),
- einen OCR-Prozessor (130) zur automatischen Auswertung von
15 Adreßinformationen aufweisenden Abbildern der SendungsOber-
flächen,
- eine Vorrichtung zur Videocodierung der die Adreßinforma-
tionen enthaltenen Abbilder der SendungsOberflächen mit min-
destens einem Videocodierplatz (200),
- einen Imagecontroller (170) zur Steuerung des Datenflusses
20 zwischen dem OCR-Prozessor (130) und der Vorrichtung zur Vi-
deocodierung aufweist, d a d u r c h g e k e n n z e i c h -
n e t, daß der Imagecontroller (170) derart ausgebildet ist,
daß das Abbild jeder innerhalb einer bestimmten Zeit mittels
Videocodierung nicht eindeutig erkannten vollständigen
25 Adreßinformation mit den bei der Videocodierung erhaltenen
Informationen über erkannte Adreßteile dem OCR-Prozessor
(130) zu einer weiteren automatischen Auswertung zur Adreßin-
terpretation zugeführt wird.

30 14. Vorrichtung nach Anspruch 13, d a d u r c h g e -
k e n n z e i c h n e t, daß eine Vorrichtung vorgesehen ist,
um Sendungen, deren Adreßinformationen nicht vollständig on-
line ausgewertet worden sind, mit einer Identifizierungsin-
formation (TID) zu versehen.

Zusammenfassung

Verfahren zur Verarbeitung von Postsachen

- 5 Bei einem Verfahren zur Bearbeitung von Sendungen in einem automatischen Adreßlesesystem, wobei ein Abbild der Adreßinformation aufweisenden Oberfläche jeder Sendung gewonnen und einer ersten automatischen Auswertung zugeführt wird, und nicht vollständig ausgewertete Adreßinformation einer ersten
- 10 Videocodierung zur Auswertung zugeführt wird, ist vorgesehen, daß die Adreßinformation derjenigen Abbilder, die bei der Videocodierung nicht vollständig ausgewertet worden sind, unter Verwendung der Ergebnisse der Videocodierung einer weiteren automatischen Auswertung zugeführt werden.

15

FIG 1